A Bayesian Approach for Uncertainty Quantification of Extreme Precipitation Projections Including Climate Model Interdependency and Nonstationary Bias

Sunyer Pinya, Maria Antonia; Madsen, Henrik; Rosbjerg, Dan; Arnbjerg-Nielsen, Karsten

Published in:
Journal of Climate

Link to article, DOI:
10.1175/JCLI-D-13-00589.1

Publication date:
2014

Document Version
Peer reviewed version

A Bayesian approach for uncertainty quantification of extreme precipitation projections including climate model interdependency and non-stationary bias

Maria Antonia Sunyer

Department of Environmental Engineering, Technical University of Denmark

Miljøvej building 113, Kgs. Lyngby 2800, Denmark.

masu@env.dtu.dk

Henrik Madsen

DHI

Dan Rosbjerg

Department of Environmental Engineering, Technical University of Denmark

Karsten Arnbjerg-Nielsen

Department of Environmental Engineering, Technical University of Denmark

Abstract

Climate change impact studies are subject to numerous uncertainties and assumptions. One of the main sources of uncertainty arises from the interpretation of climate model projections. Probabilistic procedures based on multi-model ensembles have been suggested in the literature to quantify this source of uncertainty. However, the interpretation of multi-model ensembles remains challenging. Several assumptions are often required in the uncertainty quantification of climate model projections. For example, most methods often assume that the climate models are independent or/and that changes in climate model biases are negligible. This study develops a Bayesian framework that accounts for model dependencies and changes in model biases and compares it to estimates calculated based on a frequentist approach. The Bayesian framework is used to investigate the effects of the two assumptions on the uncertainty quantification of extreme precipitation projections over Denmark. An ensemble of regional climate models from the ENSEMBLES project is used for this purpose.

The results confirm that the climate models cannot be considered independent and show that the bias depends on the value of precipitation. This has an influence on the results of the uncertainty quantification. Both the mean and spread of the change in extreme precipitation depends on both assumptions. If the models are assumed independent and the bias constant, the results will be overconfident and may be treated as more precise than they really are. This study highlights the importance of investigating the underlying assumptions in climate change impact studies, as these may have serious consequences for the design of climate change adaptation strategies.

1. Introduction

Information on the expected changes in extreme precipitation under climate change conditions is needed for the design of infrastructure such as dams, bridges and urban drainage. However, these changes are subject to numerous uncertainties. Uncertainties are introduced at the different steps in climate change impact studies, i.e. the choice of emission scenario and Global Circulation Model (GCM) as well as dynamical and/or statistical downscaling method. It is generally recognized that robust decision-making on climate change adaptation should be based on probabilistic climate projections that include these sources of uncertainty (Fowler et al. 2007; Tebaldi and Knutti 2007).

Recent studies have addressed and compared the contribution of the various sources of uncertainty (e.g. Wilby and Harris 2006; Déqué et al. 2007; Dessai and Hulme 2007; Hawkins and Sutton 2011; Räisänen and Räty 2012). In most cases the authors found, that the inherent uncertainty in climate models exceeds the uncertainty due to the natural variability, emission scenario, and statistical downscaling. Nonetheless, the relevance of each uncertainty source varies depending on the projection horizon. Natural variability and emission scenario become more important when analysing, respectively, short and long projection horizons (Hawkins and Sutton 2011).

Uncertainty in climate models is often assessed based on a multi-model framework. Large multi-model ensemble datasets have been produced in recent years. One of the largest datasets of GCMs is the ensemble recently created in the 5th Coupled Model Intercomparison Project (CMIP5) (Taylor et al. 2012). In Europe, the PRUDENCE project (Christensen et al. 2007) aimed at addressing the uncertainties in a multi-model ensemble of Regional Climate Models (RCMs). This work was continued in the ENSEMBLES project (van der Linden and Mitchell 2009). As part of the ENSEMBLES project a large dataset of state-of-the-art RCMs was created and made freely available. Currently, in the CORDEX project (http://wcrp-cordex.ipsl.jussieu.fr/) several ensembles of RCMs covering the majority of populated areas in the world are being created.

Multi-model ensembles provide valuable information for estimating the uncertainty in climate model projections. However, there are still remaining challenges in their interpretation (Knutti et al. 2010). Some of the main challenges are: the limited number of climate models in the ensemble, the interdependency of climate models, and the lack of consensus on how to evaluate model performance. Therefore, several assumptions are often needed in climate change studies (Fischer et al. 2012). For example, most studies assume that the climate models in the ensemble are independent. This assumption may not be valid since some models share similar or even identical parameterization schemes as well as submodels with identical code (Masson and Knutti 2011). Additionally, in the case of ensembles of RCMs, some models are driven by the same GCM boundary conditions. Another important assumption in climate change studies is the assumption of constant bias in the climate models. The bias is assumed to be stationary in time, i.e. its change is considered negligible. This assumption is often used in statistical downscaling and bias correction methods. However, a recent study by Boberg and Christensen (2012) showed that the bias of temperature is dependent on the value of the temperature.

In recent years, several probabilistic methods based on multi-model ensembles have been suggested in the literature, e.g. the Reliability Ensemble Averaging method suggested by Giorgi and Mearns (2002) and the Bayesian methods suggested by Tebaldi et al. (2004; 2005), Leith and Chandler (2010), and Buser et al. (2009; 2010). These methods account for some of the challenges in interpreting multi-model ensembles. The assumption of model independency is discussed in several studies, e.g. Furrer et al. (2007), Tebaldi and Knutti (2007), Buser et al. (2009; 2010), Knutti et al. (2010) and Maulet et al. (2012). Although most of these studies discussed the possible invalidity of this assumption, none of the methods include the model interdependency in the uncertainty quantification approach. Buser et al. (2009; 2010) addressed the effects of assuming constant bias. They tested the assumption of constant bias versus the assumption of temperature dependent bias

and found that the choice of constant or dependent bias assumption largely influenced the results obtained for future summer mean temperatures.

This study develops a Bayesian framework that accounts for the interdependency and the changes in the biases of the climate models. The main goal of the study is to develop and analyse a procedure for the incorporation of these two factors in the uncertainty quantification of climate model projections. The study addresses the uncertainty in extreme precipitation projections over Denmark using an ensemble of RCMs. The focus is on extreme precipitation as it is likely to be one of the most important impacts of climate change in cities (Fowler and Hennessy 1995; IPCC 2012). Hence, a quantitative assessment of the uncertainty in these projections is needed for the design of climate change adaptation in cities. Most studies have focused on the uncertainty in mean precipitation projections at rather large scales (e.g. Giorgi and Mearns, 2002; Fischer et al. 2012), and only a few studies have assessed the uncertainties in extreme precipitation projections at a regional scale (e.g. Frei et al. 2006; Fowler and Ekström, 2009).

The next section describes the case study and the data used, followed by the methodology section. Section 4 presents and discusses the results, and section 5 summarizes the findings and conclusions of the study.

2. Data and case study

The index selected to represent extreme precipitation in this study is the 95th percentile of precipitation amount on wet days, hereafter referred to as $RR_{wn}95$. Wet days are defined as days with precipitation higher or equal to 1 mm (Peterson 2005). $RR_{wn}95$ is included in the list of extreme precipitation indices defined by STARDEX (Haylock and Goodess 2004). Additionally, it is often used in climate change impact studies focusing on extreme precipitation (e.g. Beldring et al.

2008; Benestad 2010). In Denmark, under present climate conditions, $RR_{wn}95$ is approximately 13 mm day^{-1} in winter and 19 mm day^{-1} in summer.

The observational dataset used in this study is the gridded precipitation product created by the Danish Meteorological Institute known as Climate Grid Denmark (CGD). This gridded dataset has a spatial resolution of 10 x 10 km and it is based on approximately 300 stations (Scharling 1999). Daily precipitation data is available for the period 1989-2010.

The climate model output used is daily precipitation from an ensemble of 15 RCMs from the ENSEMBLES project (van der Linden and Mitchell 2009). These 15 RCMs are all the models from ENSEMBLES available at a temporal resolution of 0.22° (approximately 25 km) and covering the time period 1950-2100. This ensemble contains 11 different RCMs driven by 6 different GCMs (see Table 1). Thirteen RCMs use the same rotated pole grid system, while two models (RM5.1 and RegCM) use a Lambert conformal grid system. Daily precipitation data from the RCMs for the time periods 1989-2010 and 2081-2100 for all the 66 land grid points covering Denmark are used in this study. The future period considered here is the same as the one used in the Special Report on Managing the Risks of Extreme Events and Disasters to Advance Climate Change Adaptation (IPCC 2012).

In order to be able to compare the indices obtained from the observational dataset and the RCMs, the values of $RR_{wn}95$ estimated for each of the grids in the CGD have been re-interpolated to the 0.22° rotated pole grid system. The method used for this purpose is the natural neighbour interpolation method suggested by Sibson (1980; 1981). Additionally, and for the same reason, the two models using the Lambert conformal grid system have been re-interpolated to the rotated pole grid system using the same interpolation method.

TABLE 1 SHOULD GO HERE

3. Methodology

The methodology followed in this study is a Bayesian framework based on the approach presented by Tebaldi et al. (2004; 2005). They suggested a Bayesian statistical model that combines information from a multi-model ensemble and observations. The statistical model was used to determine probability distributions of future temperature change for 22 land regions. Here, the approach has been modified to include the interdependency and the change in bias of the climate models. A frequentist approach could have also been used (after a re-parameterization of the model) to estimate the uncertainty in the parameters of the statistical model defined here. The discussion on frequentist versus Bayesian frameworks is generally recognized and on-going (e.g. Wilks 2006; Beven 2010). One important difference between the two frameworks is in the interpretation of probability. In the frequentist view, probability represents the likelihood of an event that would be found if it was possible to take a large number of samples (Beven 2010). In the Bayesian view, probability represents the degree of belief on the occurrence of an uncertain event (Wilks 2006). According to Tebaldi et al. (2005) Bayesian methods are a natural way to represent uncertainty in the context of climate change projections. We agree with this view, and we interpret the probabilities found here as a degree of belief given the limited amount of data. Therefore, and because our statistical model is mainly inspired by Tebaldi et al. (2004; 2005), we apply a Bayesian approach to estimate the uncertainty in the model parameters. Nonetheless, we compare the results of the Bayesian analysis with a frequentist approach. The estimation of the parameters using the frequentist approach is described in Appendix B.

This section introduces the Bayesian statistical model as well as the methods used to estimate the interdependency and change in bias of the climate models. The methods described are applied to the index $RR_{wn}95$ separately for winter (December to February) and summer (June to August) season. The analysis is also carried out considering the whole year; this is referred to as annual

period in the results section. For clarity, the subscript to indicate seasons is not included in the mathematical expressions.

a. Bayesian statistical model

In a Bayesian framework probability distributions are explicitly used for quantifying the uncertainty in parameters. In this study we apply Bayesian inference, where statistical interpretations about a set of parameters, $\boldsymbol{\Theta}$, are made in terms of the conditional probability on the data, $\mathbf{D}$. Following Bayes' rule the conditional probability of $\boldsymbol{\Theta}$ depending on $\mathbf{D}$ can be expressed as:

$$p(\boldsymbol{\Theta}|\mathbf{D}) \propto p(\boldsymbol{\Theta})p(\mathbf{D}|\boldsymbol{\Theta}) \tag{1}$$

where $p(\boldsymbol{\Theta}|\mathbf{D})$ is the posterior distribution, $p(\boldsymbol{\Theta})$ is the prior distribution and $p(\mathbf{D}|\boldsymbol{\Theta})$ is the likelihood function (or sampling distribution).

For the purpose of our study here, $\boldsymbol{\Theta}$ is the set of parameters of a statistical model. As in Tebaldi et al. (2004; 2005), the statistical model is constructed based on the assumption that the values of $RR_{wn}95$ estimated from the observations and the climate models are normally distributed. The main difference between the approach suggested here and the approach by Tebaldi et al. (2004; 2005) is that the climate models are assumed to follow a multivariate normal distribution instead of independent univariate normal distributions. The use of a multivariate normal distribution allows us to take into account the interdependency between the climate models.

A multivariate normal distribution is defined for the climate models for both present and future time periods. The mean of the multivariate normal distributions is considered to be the "true mean" plus the common bias of the climate models. Additionally, a univariate distribution is defined for the observations with mean equal to the "true mean" and a known variance. Then, the statistical model for $RR_{wn}95$ can be expressed as:

$$\begin{aligned}\mathbf{X} &\sim N_M\left((\mu+\beta)\mathbf{1}, \lambda^{-1}\mathbf{R}\right)\\ \mathbf{Y} &\sim N_M\left((\nu+\alpha\beta)\mathbf{1}, (\theta\lambda)^{-1}\mathbf{R}\right)\\ X_{Obs} &\sim N\left(\mu, \sigma_{Obs}^2\right)\end{aligned} \quad (2)$$

where M is the total number of climate models; $\mathbf{X}$ and $\mathbf{Y}$ are vectors of length M containing the values of the $RR_{wn}95$ X_m and Y_m for all the climate models m = 1, ..., M for present and future conditions, respectively; X_{Obs} is the value of $RR_{wn}95$ estimated from the observations; μ and ν are scalars that represent the "true mean" of $RR_{wn}95$ for present and future conditions; β and λ^{-1} are scalars representing, respectively, the common bias and variance of the climate models for present conditions; α and θ are introduced to account for the fact that the bias and the variance might change from present to future; $\mathbf{R}$ is the correlation matrix of the climate models assumed to be constant from present to future; σ_{Obs}^2 is the variance of the observations. X_m, Y_m and X_{Obs} are estimated as the average of $RR_{wn}95$ over all grid points in the study area. In the following, $\mathbf{R}$, α, and σ_{Obs} are considered as known. Procedures for assessing these three terms will be given in the following sections. From Eq. (2) it follows that the likelihood of the data given a set of parameter values can be estimated as:

$$p\left(\mathbf{D}|\mathbf{\Theta}\right) = f_M\left(\mathbf{X}; (\mu+\beta)\mathbf{1}, \lambda^{-1}\mathbf{R}\right) f_M\left(\mathbf{Y}; (\nu+\alpha\beta)\mathbf{1}, (\theta\lambda)^{-1}\mathbf{R}\right) f\left(X_{Obs}; \mu, \sigma_{Obs}^2\right) \quad (3)$$

where f_M is a multivariate normal probability density function with dimension M, and f is a univariate normal probability density function;.

The prior distribution, $p(\mathbf{\Theta})$, can be expressed as the product of the marginal prior distributions of all the parameters. For the parameters related to the mean values (μ, ν, and β) the prior distribution is assumed to be a univariate normal distribution, while a gamma distribution is assumed for the parameters related to the variances (λ and θ). This selection of prior distributions is based on the natural conjugate prior families, i.e. the distributional forms for the likelihood and the prior are

conjugate. Table 2 summarizes the distribution and distribution parameters (hyperparameters) chosen for each of the parameters. The values of the hyperparameters have been determined to ensure that marginal prior distributions carry little information. This implies that the data available are the main factor defining the posterior distribution. Uninformative priors are chosen because we assume that "a priori" we do not have any knowledge on the distribution of the parameters. Hence, the parameters chosen for the priors should not affect the results, as long as the prior distributions are uninformative. The selection of the distributions and their parameters is inspired by the prior distributions chosen by Tebaldi et al. (2004; 2005) and Buser et al. (2009; 2010).

TABLE 2 SHOULD GO HERE

A Markov Chain Monte Carlo (MCMC) algorithm is applied using Gibbs sampling to infer the posterior distribution (Gelman et al. 2003). In the Gibbs sampler the value of each parameter is estimated conditional on the value of all the other parameters using the marginal conditional probabilities. The conditional probability distributions are shown in Appendix A. A *burn-in* period has been used in the MCMC to account for the effect of the starting distribution. As suggested in Gelman et al. (2003) the first half of the iterations has been discarded. Additionally, the sequence has been *thinned* by keeping only one in 15 iterations to avoid dependency within the MCMC. The independency of the sequence and the convergence of the algorithm have been assessed following the guidelines in Gelman et al. (2003). For this purpose, 5 sequences of length 100,000 iterations (after the *burn-in* period) have been considered.

The approach defined here necessarily relies on a number of subjective assumptions. Two simplifications considered here contrast with the approaches suggested in the literature, e.g. Tebaldi et al. (2004; 2005) and Buser et al. (2009; 2010). These are: the use of the same variance and bias for all the climate models.

The use of the same variance has been applied for two main reasons. Firstly, the focus of this study is not to estimate the reliability or weighting of the individual climate models, which would require considering different variances as suggested by Tebaldi et al. (2004; 2005). Currently, there is no consensus on the best approach to weight climate models (Knutti et al. 2010; Weigel et al. 2010), and some of the approaches suggested lead to virtually equal weights, e.g. the approach suggested in the ENSEMBLES project (Christensen et al. 2010). Secondly, by considering the same variance for all climate models it is possible to derive the conditional probability for each parameter, which allows us to use Gibbs sampling.

The bias in the climate models is assumed to be the sum of a common bias and an individual bias, i.e. for the present period the common bias is β and the individual bias of the model m is $X_m - \beta - \mu$. Along with the common bias, the same change in bias α is used for all the climate models. This differs from the individual biases considered in Buser et al. (2009; 2010). Several studies have shown that climate models have common biases in the simulation of precipitation (e.g. Boberg, et al. 2010; Sunyer et al. 2013a). A drawback of using a common bias is that the individual biases for the present and future period are assumed independent. This assumption is questionable, but comparable to the simple model defined in Tebaldi et al. (2004). Here we use this simple model to illustrate the effects of the interdependency of RCMs and change in bias in the uncertainty quantification, but further work could focus on identifying such a bias and testing its importance.

The use of common variance and bias reduces the number of parameters significantly. This allows better identification of model parameters considering the available data. If different parameters were used for each climate model, the number of parameters would be $2(M\text{-}1)$ higher than in the statistical model described above, i.e. the statistical model would have 33 parameters instead of 5 parameters. It is not clear how these simplifications affect the results. A possible effect of using a common bias is that if the RCM biases are very different, the uncertainty in β will be large. This

will affect the uncertainty in μ and v, probably by increasing their uncertainty. A similar effect could occur when applying the same variance in all the models.

b. Variables considered known in the Bayesian statistical model

1) Observed variance

The variance of $RR_{wn}95$ estimated from the observations, σ_{Obs}^2, is considered known in the Bayesian model. Since $RR_{wn}95$ is given by the regional average of the indices calculated for each grid point, the variance σ_{Obs}^2 can be estimated as the average of the variances for the individual grid points divided by the total number of independent grid points in the region. For each grid point i, σ_i^2 is calculated by applying the jackknife resampling technique (Miller, 1974; Wilks, 2006), where the index is recomputed n times using a sub-set of the data which excludes one year of data each time (n is the number of years in the sample).

As described in Bretherton et al. (1999), the number of independent grid points in the region can be estimated by dividing the total number of grid points with the decorrelation length defined as the distance between independent observations. It can be estimated using semivariograms as the distance at which the semivariance levels off. Here, we use the semivariograms and the methodology used in Sunyer et al. (2013a,b) to estimate the decorrelation length of the observational dataset.

It must be noted that here the observed variance, σ_i^2, is considered to represent the sampling error in the estimation of $RR_{wn}95$ rather than a measure of natural variability as in Tebaldi et al. (2004; 2005). Natural variability is not explicitly included in the model as done in Buser et al. (2009; 2010). The model defined in Eq. (2) considers two stationary time slices, control and future period, i.e. the value of $RR_{wn}95$ in one time slice does not depend on time. Decadal variability is also not included in the analysis and cannot be estimated with the data available. Longer observation records

and a reformulation of the model in Eq. (2) would be required to be able to introduce decadal variability in the analysis. Different ways to define the observed variance exist and one must be aware that the choice of model may influence the results.

2) Interdependency of RCMs

As described in the previous section, the interdependency of the climate models is included in the Bayesian approach through the correlation matrix of the climate models, **R**. This is estimated using the approach suggested by Pennell and Reichler (2011) to assess the amount of independent information in an ensemble of climate models. The approach is based on the idea that two models are not independent because they lead to different results but because they follow different paths to reach the results. Hence, according to Pennell and Reichler (2011), a suitable approach for assessing the independency of the climate models is to statistically analyse their errors.

A metric, *d*, that represents the error of the climate models for present conditions is used to estimate **R**. The metric is calculated separately for all the grid points of each climate model. It is estimated from the values of the individual model error and the ensemble average error, which again are estimated from the value of $RR_{wn}95$ obtained from the climate models and the observations for the present time period. The individual model error is estimated by subtracting the observed values from the climate model values, and dividing this difference by the standard deviation of the observations. The standard deviation of the observations at each grid point is estimated as the interannual variability. The ensemble average error is estimated as the average of the individual model errors. The metric is then calculated by removing from the individual model error the part present in the ensemble average error. This is done using the correlation, *r*, between the individual model errors and the ensemble average error. For each grid point, *i*, and climate model, *m*, the metric, $d_{i,m}$, is estimated as: $d_{i,m} = e^*_{i,m} - r_m \bar{e}^*_i$, where $e^*_{i,m}$ and $\bar{e}^*_i$ are the standardized values of the

individual model error and the ensemble average error, respectively. More details on the method used to estimate d can be found in Pennell and Reichler (2011) and Sunyer et al. (2013a).

The matrix $\mathbf{R}$ is then estimated from the values of d obtained for each climate model. The elements of $\mathbf{R}$ are the correlation coefficients corr($\mathbf{d}_{\cdot,m}$, $\mathbf{d}_{\cdot,n}$), where $\mathbf{d}_{\cdot,m}$ and $\mathbf{d}_{\cdot,n}$ are the vectors containing the value for all the grid points of the models m and n, respectively. In the Bayesian approach the absolute values of the matrix $\mathbf{R}$ are used in the multivariate normal distribution for both present and future. Hence, it is assumed that a negative correlation also implies that the models are interdependent.

3) Change in bias

In the statistical model defined in Eq. (2), the possible change in bias is accounted for by the term α. Boberg and Christensen (2012) suggested a procedure that uses the area-averaged monthly bias from the RCMs driven by re-analysis data to study biases of RCM simulations. Here we suggest to modify the procedure and use the value of $RR_{wn}95$ estimated by the RCMs driven by GCMs and compare with observations at each grid point to calculate the bias, because the main objective is to study the biases of the projected values of $RR_{wn}95$. The procedure is based on two main steps. First, for each grid point, i, and RCM, m, the bias, $ABias_{m,i}$, is estimated as $X_{m,i}$-$X_{Obs,i}$. Then, a linear regression is estimated between $\mathbf{ABias}_{\cdot,\cdot}$ and $\mathbf{X}_{\mathbf{Obs},\cdot}$, i.e. all the climate models are used to estimate the linear regression. It must be noted that this approach ignores the fact that the grid points are correlated. This is expected to have minor influence on the linear regression.

The slope parameter of the linear regression, A, is subsequently used to estimate the change in the bias, α. Linear regressions are derived for the two seasons and annual data, i.e. a different value of α is used for winter, summer and annual period. α depends on the constant A and the parameters β, μ, and v. α is estimated from the linear regression and Eq. (2) as:

$$\alpha = 1 + \frac{A}{\beta}(\nu - \mu) \quad (4)$$

Eq. (4) is used in the MCMC to estimate α in each iteration. It is assumed that the estimated structure of the bias over the region and the linear regression approach can be used to describe the change in bias in the future. More elaborate procedures could be developed to assess the possible change in bias under future conditions. The simple procedure presented here is considered appropriate to illustrate the effects of change in bias on the uncertainty in climate projections.

4. Results

The results of this study are presented in three parts. First, the main results regarding the interdependency and the change in bias of the RCMs are described. Second, the results from the uncertainty quantification are presented and discussed. In the last part, the effects of considering the interdependency and the bias are investigated. All the results are presented for $RR_{wn}95$ for winter, summer and annual period.

a. Interdependency and change in bias of RCMs

The interdependency of the RCMs has been analysed using the correlation matrix, **R**. The average of all the elements in **R** (without considering the diagonal) is 0.25 in winter, 0.20 in summer and 0.27 in the annual period. The minimum and maximum correlations found are 0 and 0.68 in winter, 0 and 0.74 in summer, and 0 and 0.80 in the annual period, respectively. The average values for the three periods are similar, with the average for the summer period slightly lower. The maximum value found is slightly higher in the annual period. The maximum correlations found for each period correspond to different RCM-GCM combinations. In winter the highest correlation is found for two RCM-GCMs with different RCM and GCM (RCA-HadCM3 and RACMO-ECHAM5); in summer for two RCM-GCMs driven by the same GCM (RegCM-ECHAM5 and HIRHAM-ECHAM5); and

in the annual period for two RCM-GCMs from the Hadley centre (HadRM3Q16-HadCM3Q16 and HadRM3Q0-HadCM3Q0).

The pairs of RCM-GCMs leading to correlations higher than 0.6 are compared for the three periods to assess whether high correlation between the RCM-GCMs arises from the same RCM driven by different GCMs, or vice versa. In the winter period, high correlations are obtained for five pairs of RCM-GCMs, three pairs of RCM-GCMs with the same RCM but different GCMs, and two pairs of RCM-GCMs with different RCMs and GCMs. In the summer period, high correlations are found for two pairs of RCMs driven by the same GCM and for three pairs of RCM-GCMs with different RCMs and GCMs. In the annual period, high correlations are found for four pairs of RCM-GCMs with the same RCM and different GCMs, two pairs with different RCMs and GCMs, and one pair of two RCM-GCMs from the Hadley centre. In general, there are more combinations with the same RCM but different GCMs leading to high correlations than different RCMs driven by the same GCM. In addition, some RCM-GCMs using different RCMs and GCMs also lead to high correlations. This indicates that other factors than the RCM or GCM used have a large influence on the correlation between RCM-GCMs. A detailed analysis of these factors is outside the scope of this study and is not addressed further here.

The results show that the RCMs are interdependent and that there are no significant differences between the degree of interdependency for the summer, winter and annual period. It must also be noticed that in the three periods some RCMs are found to be uncorrelated. A more detailed analysis of the interdependency of this ensemble of RCMs over Denmark can be found in Sunyer et al. (2013a).

The change in bias is estimated using the bias of the RCMs over the region. Figure 1 shows the linear regression found between the bias and the observations for winter, summer and annual

period. The figure shows both the linear regression fitted for each of the RCMs individually and the linear regression fitted considering all the RCMs together. For all three periods, all the individual linear regressions show a decrease and a shift in the sign of the bias for increasing precipitation. The negative slope from the linear regression considering all the RCM-GCMs together is more pronounced in the summer period. The values of the slope, A, from the linear regression that will be used in the MCMC to estimate α are -0.54, -0.78 and -0.57 for winter, summer and annual period, respectively.

FIGURE 1 SHOULD GO HERE

For both seasons and the annual period, the correlation found between the RCMs and the bias depending on the precipitation value point towards the invalidity of assuming independency and constant bias. Hence, it is important to include these two results in the uncertainty quantification. The results show that there is a similar interdependency of the RCMs for the three periods, but that the change in bias is more pronounced for the summer period.

b. Uncertainty quantification

This section presents the probability distributions estimated for the parameters of the statistical model described in the methodology section. The convergence and autocorrelation of the outputs from the Gibbs sampling have been analysed, leading to the conclusion that convergence was reached and that the outputs are not autocorrelated (results not shown). Hence, the results from the sampling algorithm can be used to estimate the marginal and joint posterior distributions of the parameters.

First, the assumption of normal distributions in Eq. (2) is analysed. Figure 2 compares the values of **X** and **Y** with theoretical values from a normal distribution. For simplicity the interdependency of

the RCMs is not considered here. All the values of **X** and **Y** fall within the 95% confidence limits, and hence they can be considered to approximately follow a normal distribution.

FIGURE 2 SHOULD GO HERE

Figure 3 shows the marginal posterior and prior distribution for each of the parameters for winter, summer, and annual period. The values of $RR_{wn}95$ found for each of the RCMs and the observations are also shown. For all the parameters and for all three periods, the prior and the posterior distributions are noticeably different. This shows that the data inputs to the model (**X**, **Y**, X_{Obs}, σ_{Obs}, **R**, and A) exert the main influence on the posterior distribution.

For the winter period, X_{Obs} lies well within the range of the values from the RCMs, while for the summer and annual periods most of the RCMs have smaller values than the observations. The larger difference between RCMs and observations found for the summer period is likely to be due to the nature of extreme events in summer, which are mainly caused by convective precipitation. RCMs are better at representing extreme events caused by frontal precipitation than convective precipitation (e.g. Herrera et al. 2010; Fowler et al. 2005). In agreement with the results found here, several studies have shown that RCMs have larger biases and inter-model differences in the summer period (Frei et al., 2006; Kendon et al. 2008; Fowler and Ekström, 2009).

For the three periods, the posterior distributions of μ are sharp compared to the individual values from the RCMs. This is because the standard deviation estimated for the observations is considerably smaller than the spread of the values from the RCMs. For the summer and annual period this is also due to the influence of the bias in some RCMs. σ_{Obs} is 0.20, 0.30 and 0.15 mm day^{-1} for winter, summer and annual period, respectively. Small values of σ_{Obs} lead to less uncertainty in the parameter μ. Conversely, all the values of the RCMs are encompassed in the posterior distribution of v. This is a combined effect of the fact that the value of σ_{Obs} does not affect

this parameter and that all the RCMs have the same weight. It must also be noted that the median of the distributions increases from present to future for both winter and annual periods, but decreases for the summer period.

The bias found for the RCMs is centred on approximately 0 mm day^{-1} for the winter period (the median is -0.01 mm day^{-1}). For the summer and annual period, the medians of the bias are -2.25 and -1.86 mm day^{-1}, respectively. The uncertainty in this parameter is larger for the summer and annual period than for the winter period. The values of the 5^{th} and 95^{th} percentiles are -1.28 and 1.14 mm day^{-1} for the winter period, -4.68 and 0.40 mm day^{-1} for the summer period, and -3.74 and 0.16 mm day^{-1} for the annual period. This shows that the RCMs have a tendency towards underestimation of extreme precipitation for both the summer and the annual period. This is likely due to the underestimation of extreme events caused by convective precipitation as discussed above. The distributions for these two periods resemble bimodal distributions, where one mode corresponds to a negative bias and the other is approximately at 0 mm day^{-1}. This is likely due to the fact that the distribution of the bias is dependent on both the values found for present and future climate. For the present climate the bias found is negative, while in the future climate due to the lack of observations the bias tends to 0 mm day^{-1}.

The distribution of λ, which represents the reliability (the inverse of the variance) of the RCMs for present climate, varies depending on the period. Both the median and the spread are higher for winter than for summer and annual period. This reflects a lower reliability of the RCMs for these two periods. The lowest values are found for the summer period. This is due to the larger differences between the RCMs for this period. The median of the posterior distribution of λ for the winter period is 0.63 mm^{-2} day^{2}. This corresponds to a standard deviation of 1.27 mm day^{-1}, which is considerably higher than the estimated standard deviation of the observational dataset. The same is observed for the summer and annual period, where the median of λ corresponds to a standard

deviation of 3.03 mm day^{-1} and 2.32 mm day^{-1}, respectively. These results are in agreement with the general perception that the reliability of the observations should be higher than for the RCMs.

The distribution of θ, which represents the change in the reliability from present to future climate, also varies depending on the period considered. The median for the winter period is 1.02, pointing to a small change in the reliability. However, the values range from a decrease to an increase in the reliability (5^{th} percentile of 0.41 and a 95^{th} percentile of 2.65). As in the case of λ, lower values and lower uncertainty are found for the summer and annual period. The medians are 0.41 and 0.44, the 5^{th} percentiles 0.17 and 0.18, and the 95^{th} percentiles 1.03 and 1.10 for the summer and the annual period, respectively. For these two periods, most of the values are lower than one. This is caused by the larger inter-model differences in the future period and shows the decrease in the reliability of the RCMs.

Figure 3 also shows the mean of the parameters obtained using the Bayesian and frequentist approach; the values for all the parameters are virtually the same except for β in winter and θ. A more detailed comparison and description of the differences between the two approaches can be found in Appendix B.

FIGURE 3 SHOULD GO HERE

The marginal posterior distributions of the parameters provide information about the uncertainty and the expected value of the parameters. Moreover, the joint posterior distribution of the parameters provides information about the relation between the parameters. Figure 4 shows the histogram for each of the parameters and the scatter plot of all the combinations of pairs of parameters for the summer period. It also shows the correlation coefficient for each pair of parameters.

The mean values for present and future conditions are practically uncorrelated. This is due to the small contribution of the RCMs in the distribution of μ compared to observations, and differs from the correlation between the values for the present and future obtained from the RCMs. These two parameters (μ and v) have a negative correlation with β. This result follows both from Eq. (2) and from the analysis of the conditional probability of β (see Appendix A). The conditional probability shows that for small values of μ and v the bias will be positive (i.e. RCMs overestimate), while for large values the bias will be negative (i.e. RCMs underestimate). This is also in agreement with the values found from the analysis of change in bias.

The three parameters μ, v, and β are not correlated with the parameters λ and θ. However, it can be observed that for high values of λ the range of μ, v, and β decreases. The same relation is observed between these parameters and θ. This can be interpreted as: the uncertainty in μ, v, and β is low when the reliability in the RCMs is high (high values of λ and θ). The same interpretation can be drawn from the analysis of the conditional probabilities. The parameters λ and θ also show the expected relationship. For higher values of λ, smaller values of θ and vice versa. Similar correlations were obtained for the winter and annual period (results not shown).

When using the priors shown in Figure 3 some posterior values of v are negative, which cannot be interpreted physically. The probability of obtaining a negative value is very low (approximately 0.2%) and could be avoided by using slightly more informative priors.

FIGURE 4 SHOULD GO HERE

c. Effect of interdependency and change in bias

The previous section analyses the uncertainty and joint distribution of the parameters in the statistical model. This section focuses on the effects of taking into account the interdependency and change in bias of the RCMs. A set of tests using different assumptions have been defined for this

purpose. First, three tests considering either independency or interdependency and change in bias or constant bias are carried out. Then, a total of 45 tests are run using different levels of interdependency and change in bias of the RCMs. The main output analysed from these tests is the change in $RR_{wn}95$ from present to future. This change, referred to as change factor (CF), is defined as the relative change between μ and v, i.e. CF = v/μ. The value of CF has been estimated in each of the iterations of the MCMC runs from which the distribution of CF has been derived.

Figure 5 shows the results from the first three tests: (1) the RCMs are independent and the bias constant, (2) the RCMs are interdependent and the bias constant, (3) the RCMs are interdependent and the bias changes from present to future. This last test is the one for which the results have been analysed in detail in the previous section. In the test assuming that the RCMs are independent, **R** is set equal to the identity matrix. Similarly, in the tests where the bias is assumed constant A is set to 0 (i.e. α equal to 1). Figure 5 shows the marginal posterior distributions for μ, v and, CF for the three different tests for winter, summer, and annual period.

For μ, the posterior distributions obtained for the three tests are similar. The median as well as the 5^{th} and 95^{th} percentiles for the three tests differ less than 0.1 mm day^{-1} for all the periods. Larger differences are found in the distributions of v. The uncertainty in v found for the test assuming independency is smaller than for the other tests. This is due to the fact that less independent information is considered in the tests assuming interdependency. This leads to a lower reliability in the RCMs, which in turn leads to higher uncertainty in v. The effects of assuming constant or changing bias are also noticed in the distribution of v. The change in bias has different effects depending on the season. In winter and annual periods, the median of v increases due to a more negative bias. The median of $\alpha\beta$ in winter and annual periods is -0.56 and -3.74 mm day^{-1}, respectively. In the summer period, the bias in the future period (the median of $\alpha\beta$ is -2.35 mm day^{-1}) is approximately equal to the bias in the present period. This leads to values of v similar to when

the bias is considered constant. Additionally, for the winter summer and annual period, the uncertainty in v increases when the change in bias is introduced. This is due to an increase in the uncertainty of the bias for the future period.

A similar difference as the one obtained for v between the three tests is observed in the results for CF. The uncertainty of this variable is higher for the tests accounting for the interdependency of the RCMs for all the periods. In addition, the uncertainty is also higher for the test accounting for change in bias. The median of CF depends on the test. In all seasons, the median of CF in the test assuming independency is similar to the median estimated directly from the RCMs (for climate model m, CF_m is estimated as Y_m/X_m). In the winter period, a larger median is obtained when the change in bias and interdependency are taken into account. For this season, large differences are also observed for the 5^{th} and 95^{th} percentiles being: 1.11 and 1.21 for test (1); 1.02 and 1.26 for test (2); and 1.01 and 1.47 for test (3). Even though the uncertainty in the tests where interdependency is taken into account is larger, all tests show an increase of extreme precipitation for winter at a 5% level of significance.

For the summer period, similar differences as in v are obtained in CF. As in the winter period, the comparison of the 5^{th} and 95^{th} percentiles points out the differences in the uncertainty of CF being: 0.98 and 1.13 for test (1); 0.78 and 1.25 for test (2); and 0.60 and 1.40 for test (3). For this season, the tests do not agree on the direction of the change. Similar results are obtained for the annual period regarding the uncertainty in CF. The values of the median found for this period are 1.09, 1.10 and 1.22 for test (1), (2), and (3), respectively. The 5^{th} and 95^{th} percentile for tests (2) and (3) vary between a decrease and an increase of extreme precipitation, while the values for test (1) point to an increase of extreme precipitation. In summary, for the summer and annual period, even though most of the RCMs point towards an increase in extreme precipitation, the interdependency between

1 the RCMs means that (given the information available) a hypothesis of a decrease in extreme
2 precipitation cannot be rejected on the 5% significance level.

3 FIGURE 5 SHOULD GO HERE

4 Appendix B compares the standard errors of μ, v, and CF and describes the differences between the
5 two approaches.

6 The difference obtained for the three tests can be further analysed by considering the joint
7 distributions of μ and v. Figure 6 compares the scatter plots obtained for each of the tests. The
8 correlation between the two parameters is higher for the test where the RCMs are assumed
9 independent. For example, for the summer the correlation coefficient is 0.32 for test (1) and 0.1 for
10 test (2) and (3). In other words, for a specific value of μ the range of possible values of v is larger in
11 the tests where the interdependency of the RCMs is taken into account. This contributes to the
12 higher uncertainty of CF found for these tests. Figure 6 also highlights the differences in the
13 uncertainty for winter, summer and annual period. As shown in Figure 5, for all three tests, the
14 range of values of v in the winter period is considerably smaller than the range of values obtained
15 for summer and annual period. In the case of μ, the range of values obtained for the annual period is
16 smaller than for the winter and summer period. This is due to the lower variance of the observations
17 for this period. Additionally, the correlation between μ and v is higher for the three tests for the
18 winter period. This shows that, in agreement with the results found in Figure 3 and Figure 5, there is
19 a lower uncertainty in the winter of CF.

20 FIGURE 6 SHOULD GO HERE

21 In order to study in more detail the influence of the interdependency and the changes in the bias, 45
22 MCMC simulations have been run using different values of **R** and A. In these simulations, artificial
23 values of **R** and A are used. In **R**, the same correlation coefficients are used for all the pairs of

RCMs ranging from 0 to 0.8 (all the values in the diagonal are kept equal to 1). The values of A are set to range between -0.8 to 0.8. The mean and the standard deviation of CF are analysed for each of these simulations. Due to the use of artificial values of $\mathbf{R}$ and A, the results of these 45 simulations cannot be directly compared to the results of the three tests described above, except in the case of independency and constant bias (test 1).

Figure 7 summarizes the results of the 45 MCMC simulations for the annual period. The y- and x-axes display the values used to define $\mathbf{R}$ and A, respectively. The results show that the uncertainty in CF is influenced by both $\mathbf{R}$ and A. The standard deviation of CF increases for higher values of $\mathbf{R}$ and it decreases for smaller absolute values of A. This is in agreement with the results from the three tests. The mean CF is also influenced by both $\mathbf{R}$ and A, but in this case the main influence is the value of A. The mean CF decreases for increasing values of A ranging from -0.8 to approximately 0.4, and it increases for increasing values of A higher than 0.4. The mean values of CF range from virtually no increase (mean CF equal to 1.03) to an increase of 40%, and the standard deviation varies from 0.03 to 0.86. The two smallest standard deviations are obtained for the cases where the models are considered independent and A equal to 0.2 and 0 (test (1)), while the largest standard deviation is obtained for correlations of 0.8 and A equal to -0.8.

FIGURE 7 SHOULD GO HERE

5. Conclusions

A Bayesian approach based on a multi-model ensemble of RCMs has been developed to assess the uncertainty in extreme precipitation projections over Denmark. This approach accounts for both interdependency and linear changes in the bias of the RCMs. The results of the Bayesian approach are compared to estimates calculated based on a frequentist approach.

The analyses of the outputs from the RCMs show that, for the region of study, both the assumption of independency and constant bias are unlikely to be valid. Hence, interdependency and change in bias are included in the uncertainty estimation. The uncertainty estimated for $RR_{wn}95$ is higher for summer than for winter both for the present and future time period. This is due to the higher reliability in the values of the RCMs for winter and a lower observed variance. The lower reliability in the summer period is explained by the larger differences between the RCMs and between the RCMs and the observations. For the annual period, the uncertainty for the present is lower than for the winter period, but it is higher for the future period. The change estimated from present to future is more uncertain for the summer and annual period than for the winter. The change in the distribution median points towards an increase in extreme precipitation for winter and annual period, and to approximately no change in extremes for the summer period.

The influence of accounting for the interdependency and the change in bias of the RCMs has been tested. These two assumptions do not affect significantly the distribution of the mean of $RR_{wn}95$ for the present. However, they largely affect the distribution for the future period, which again influences the distribution of the change factor. A higher uncertainty in the change factor is obtained when the interdependency and change in bias of the RCMs is taken into account.

The methods used to estimate the interdependency and the change in bias are subject to several assumptions and uncertainties. Therefore, a total of 45 MCMC simulations have been run using different (artificial) levels of interdependency and change in the bias of the RCMs. The uncertainty of the change factor increases for higher RCM correlations and it decreases for smaller absolute values of the change in bias. The mean CF is also influenced by both the interdependency and change in bias, but mainly by the change in bias.

The results of this study highlight the importance of the assumptions taken in uncertainty estimation methods for climate change impact studies. In this case study, the assumption of independency and constant bias of the climate models has led to lower uncertainty estimates of the change in extreme precipitation in comparison to the more realistic assumption of interdependency and change in bias. This highlights the risk of reaching overconfident results, which may lead to overconfident decisions regarding adaptation to climate change.

The approach presented here addresses some of the assumptions often taken in uncertainty estimation methods for climate change projections. It suggests a way in which the interdependency and change in bias can be introduced in the uncertainty quantification of climate model projections. However, there are still remaining challenges and further work is needed to address some of the limitations of this approach. The approach could be further extended by relaxing some of the assumptions such as the constant interdependency of the RCMs from present to future, independency between individual errors from present to future, the assumption of a common bias of the RCMs, or the definition of observed variance. Hence, the specific results found here should be treated with care. The main message of this study is the need to include the interdependency of climate models and change in bias in the uncertainty quantification of climate change projections.

Acknowledgments

This work was carried out with the support of the Danish Council for Strategic Research as part of the project RiskChange, contract no. 10-093894 (http://riskchange.dhigroup.com). The Climate Grid Denmark dataset is a product of the Danish Meteorological Institute. The data from the RCMs used in this work was funded by the EU FP6 Integrated Project ENSEMBLES contract no. 05539 (http://ensembles-eu.metoffice.com), whose support is gratefully acknowledged. We would like to

thank the three reviewers for their constructive comments, which have significantly contributed to improve the quality of this paper.

APPENDIX A

Full conditional probabilities

The full conditional probabilities for each of the parameters are:

$$\mu \mid \ldots \sim N\left(\widetilde{\mu}, \left(\lambda \sum_{p=1}^{M}\sum_{m=1}^{M} s_{m,p} + \sigma_{Obs}^{-2} + \sigma_{\mu}^{-2}\right)^{-1}\right);$$

$$\nu \mid \ldots \sim N\left(\widetilde{\nu}, \left(\theta\lambda \sum_{p=1}^{M}\sum_{m=1}^{M} s_{m,p} + \sigma_{\nu}^{-2}\right)^{-1}\right);$$

$$\beta \mid \ldots \sim N\left(\widetilde{\beta}, \left(\left(\lambda + \theta\lambda\alpha^{2}\right)\sum_{p=1}^{M}\sum_{m=1}^{M} s_{m,p} + \sigma_{\beta}^{-2}\right)^{-1}\right);$$

$$\lambda \mid \ldots \sim Ga\left(a_{\lambda} + M, \left(\frac{1}{2}(\mathbf{X}-\mu-\beta)\mathbf{R}^{-1}(\mathbf{X}-\mu-\beta) + \frac{1}{2}\theta(\mathbf{Y}-\nu-\alpha\beta)\mathbf{R}^{-1}(\mathbf{Y}-\nu-\alpha\beta) + \frac{1}{b_{\lambda}}\right)^{-1}\right);$$

$$\theta \mid \ldots \sim Ga\left(a_{\theta} + \frac{M}{2}, \left(\frac{1}{2}\lambda(\mathbf{Y}-\nu-\alpha\beta)\mathbf{R}^{-1}(\mathbf{Y}-\nu-\alpha\beta) + \frac{1}{b_{\theta}}\right)^{-1}\right);$$

$$\widetilde{\mu} = \frac{\frac{1}{2}\lambda\sum_{p=1}^{M}\left\{\sum_{m=1}^{M}\left[(X_m-\beta)s_{m,p}\right] + (X_p-\beta)\sum_{m=1}^{M} s_{m,p}\right\} + \frac{X_{Obs}}{\sigma_{Obs}^{2}} + \frac{\mu_0}{\sigma_{\mu}^{2}}}{\lambda\sum_{p=1}^{M}\sum_{m=1}^{M} s_{m,p} + \sigma_0^{-2} + \sigma_{\mu}^{-2}};$$

$$\widetilde{\nu} = \frac{\frac{1}{2}\theta\lambda\sum_{p=1}^{M}\left\{\sum_{m=1}^{M}\left[(Y_m-\alpha\beta)s_{m,p}\right] + (Y_p-\alpha\beta)\sum_{m=1}^{M} s_{m,p}\right\} + \frac{\nu_0}{\sigma_{\nu}^{2}}}{\theta\lambda\sum_{p=1}^{M}\sum_{m=1}^{M} s_{m,p} + \sigma_{\nu}^{-2}};$$

$$\tilde{\beta} = \frac{\frac{1}{2}\lambda\sum_{p=1}^{M}\left\{\sum_{m=1}^{M}\left[(X_m-\mu)s_{m,p}\right]+(X_p-\mu)\sum_{m=1}^{M}s_{m,p}\right\}}{(\lambda+\theta\lambda\alpha^2)\sum_{p=1}^{M}\sum_{m=1}^{M}s_{m,p}+\sigma_\beta^{-2}} +$$

1

$$+\frac{\frac{1}{2}\theta\lambda\alpha\sum_{p=1}^{M}\left\{\sum_{m=1}^{M}\left[(Y_m-\nu)s_{m,p}\right]+(Y_p-\nu)\sum_{m=1}^{M}s_{m,p}\right\}+\frac{\beta_0}{\sigma_\beta^2}}{(\lambda+\theta\lambda\alpha^2)\sum_{p=1}^{M}\sum_{m=1}^{M}s_{m,p}+\sigma_\beta^{-2}}$$

2 where a_λ and a_θ are the shape parameters of the prior distribution of λ and θ, respectively; b_λ and b_θ
3 are the scale parameters of the prior distribution of λ and θ, respectively; μ_0 and σ^2_μ are the mean
4 and variance of the prior distribution of μ; ν_0 and σ^2_ν are the mean and variance of the prior
5 distribution of ν; β_0 and σ_β^2 are the mean and variance of the prior distribution of β; $s_{m,p}$ is the
6 element [m,p] of the inverse of the correlation matrix of the RCMs, $\mathbf{R}^{-1}$. Ga(a,b) denotes a gamma
7 distribution with shape a and scale b and N(c,d^2) denotes a normal distribution with mean c and
8 variance d^2.

9

10

11

APPENDIX B

Frequentist approach

To estimate the parameters of the model defined in Eq. (2) using the frequentist approach a re-parameterization of the model is needed. The re-parameterized model considers five parameters: μ, μ_X, μ_Y, σ_X^2, and σ_Y^2 and can be expressed as:

$$
\begin{aligned}
&X_{obs} \sim N(\mu, \sigma_{obs}^2) \\
&\mathbf{X} \sim N(\mu_X \mathbf{1}, \sigma_X^2 \mathbf{R}) \\
&\mathbf{Y} \sim N(\mu_Y \mathbf{1}, \sigma_Y^2 \mathbf{R})
\end{aligned}
$$

The new set of parameters is defined using the parameters in Eq. (2) as:

$$
\begin{aligned}
&\mu_X = \mu + \beta \\
&\mu_Y = \nu + \alpha\beta = \mu + \beta + (A+1)(\nu - \mu) \\
&\sigma_X^2 = \lambda^{-1} \\
&\sigma_Y^2 = (\theta\lambda)^{-1}
\end{aligned}
$$

Following the frequentist approach the parameters in the re-parameterized model can be estimated as:

$$\hat{\mu} = X_{obs}; \quad \hat{\mu}_X = \frac{\mathbf{1}^{\mathrm{T}} \mathbf{R}^{-1} \mathbf{X}}{\mathbf{1}^{\mathrm{T}} \mathbf{R}^{-1} \mathbf{1}}; \quad \hat{\mu}_Y = \frac{\mathbf{1}^{\mathrm{T}} \mathbf{R}^{-1} \mathbf{Y}}{\mathbf{1}^{\mathrm{T}} \mathbf{R}^{-1} \mathbf{1}}$$

$$\hat{\sigma}_X^2 = \frac{(\mathbf{X} - \hat{\mu}_X \mathbf{1})^{\mathrm{T}} \mathbf{R}^{-1} (\mathbf{X} - \hat{\mu}_X \mathbf{1})}{M - 1}; \quad \hat{\sigma}_Y^2 = \frac{(\mathbf{Y} - \hat{\mu}_Y \mathbf{1})^{\mathrm{T}} \mathbf{R}^{-1} (\mathbf{Y} - \hat{\mu}_Y \mathbf{1})}{M - 1};$$

and the variance of the main estimated parameters ($\hat{\mu}$, $\hat{\mu}_X$ and $\hat{\mu}_Y$) is:

$$\mathrm{var}(\hat{\mu}) = \hat{\sigma}_{Obs}^2; \quad \mathrm{var}(\hat{\mu}_X) = \frac{\hat{\sigma}_X^2}{\mathbf{1}^{\mathrm{T}} \mathbf{R}^{-1} \mathbf{1}}; \quad \mathrm{var}(\hat{\mu}_Y) = \frac{\hat{\sigma}_Y^2}{\mathbf{1}^{\mathrm{T}} \mathbf{R}^{-1} \mathbf{1}}$$

Then, the estimates of the parameters in Eq. (2) can be obtained as:

$$\hat{\mu} = X_{obs};\ \ \hat{\nu} = \frac{\hat{\mu}_Y - \hat{\mu}_X}{A+1} + \hat{\mu};\ \ \hat{\beta} = \hat{\mu}_X - \hat{\mu};\ \ \hat{\lambda} = (\hat{\sigma}_X^2)^{-1};\ \ \hat{\theta} = \frac{\hat{\sigma}_X^2}{\hat{\sigma}_Y^2};$$

and the standard error of $\hat{\mu}$, $\hat{\nu}$, and $\hat{\beta}$:

$$\mathrm{SE}_{\hat{\mu}} = \sqrt{\sigma_{Obs}^2};\ \ \mathrm{SE}_{\hat{\nu}} = \sqrt{\frac{\sigma_X^2 + \sigma_Y^2}{(A+1)^2 \mathbf{1}^{\mathbf{T}} \mathbf{R}^{-1} \mathbf{1}} + \sigma_{Obs}^2};\ \ \mathrm{SE}_{\hat{\beta}} = \sqrt{\frac{\sigma_X^2}{\mathbf{1}^{\mathbf{T}} \mathbf{R}^{-1} \mathbf{1}} + \sigma_{Obs}^2}$$

It must be noted that the standard error of $\hat{\nu}$ is estimated assuming independency between $\hat{\mu}$, $\hat{\mu}_X$ and $\hat{\mu}_Y$, and that the standard error of $\hat{\beta}$ assumes independency between $\hat{\mu}$ and $\hat{\mu}_X$. The standard errors of the parameter estimates $\hat{\lambda}$ and $\hat{\theta}$ can be estimated from the inverse of $\hat{\sigma}_X^2$ and $\hat{\sigma}_Y^2$. From the re-parameterised model it can be found that the inverse of $\hat{\sigma}_X^2$, i.e. $\hat{\lambda}$, follows a gamma distribution with variance:

$$SE(\hat{\lambda}) = \sqrt{2(M+2)\left[(X - \mu_X \mathbf{1})^{\mathbf{T}} \mathbf{R}^{-1} (X - \mu_X \mathbf{1})\right]^{-2}}$$

The variance of the inverse of $\hat{\sigma}_Y^2$ can be estimated in a similar way. The variance of $\hat{\theta}$ can be estimated using independent draws from these distributions.

Finally, the mean of CF can be found from $\hat{\nu} / \hat{\mu}$. The standard error can be obtained using independent draws from the distributions of $\hat{\mu}$ and $\hat{\nu}$, which follow a normal distribution with mean $\hat{\mu}$ and $\hat{\nu}$ and variance $(\mathrm{SE}_{\hat{\mu}})^2$ and $(\mathrm{SE}_{\hat{\nu}})^2$, respectively.

In addition to the difference in the interpretation of probability in the Bayesian and frequentist approach, there are two other differences to be noted. First, it is not possible to estimate the probability density function of the parameters from the frequentist approach. Secondly, the

frequentist approach does not require MCMC or the use of priors which might have an influence on the results. These differences should be kept in mind when comparing the results of the two approaches.

Comparison of Bayesian and frequentist approach

The comparison of the results obtained using the frequentist and Bayesian approach helps to address the influence of the priors and MCMC in the results of the Bayesian approach. Figure B1 shows the mean and standard error estimated using the frequentist approach and the Bayesian approach for the three tests (see description of the tests in section 4.c).

FIGURE B1 SHOULD GO HERE

The mean and standard error of μ for the three tests and for the three periods are virtually the same in the frequentist and Bayesian approach. For the other parameters, the results of the mean and standard error are also similar, but slight differences are seen between the two approaches.

The largest differences between the mean of v obtained using the Bayesian and the frequentist approach is 0.02, 0.4 and 0.03 mm day^{-1} in winter, summer, and annual period, respectively. In the case of the mean of β, the largest differences are 0.16, 0.1, and 0.05 mm day^{-1}, in winter, summer, and annual period, respectively. The largest differences in the mean of v and β are obtained for test (3). A possible reason for the differences in β is the fact that in the re-parameterized model β is estimated from X_{Obs} and **X**, while in the Bayesian model in Eq. (2) β is also influenced by **Y**. The standard errors of v and β are similar for the two approaches. In some cases it is slightly larger for the Bayesian approach, and in others it is larger in the frequentist approach (e.g. v in test (3) in winter and annual period).

The mean and standard errors of λ are also very similar in the two approaches. The largest differences are seen in the mean of θ, which is smaller in the frequentist approach. The largest difference is obtained for the winter period, in this case the Bayesian approach leads to 1.22 and the frequentist approach to 1.01 (which is approximately equal to the median of θ in the Bayesian approach). For this parameter similar differences between the frequentist and Bayesian approach are obtained for the three tests.

The mean and standard error of CF are similar in the two approaches. As in the case of v, the largest difference between the two approaches is found for the mean CF for summer and test (3). In this case, the Bayesian approach leads to a mean CF of 1.02 and the frequentist approach leads to 1.04. The differences in the standard errors are similar to the ones found for v and β.

The influence of the re-parameterization has been assessed by applying the Bayesian inference technique (use of Gibbs sampling and priors shown in Table 2) defined in section 3 to the re-parameterized model. In this case, the largest difference between the Bayesian and frequentist approaches in the mean of v, β, and θ is 0.07 mm day^{-1}, 0.02 mm day^{-1}, and 0.004, respectively.

There are several reasons for which differences in the parameters arise between the two approaches. These include: the priors selected in the Bayesian approach, the convergence of the MCMC in the Bayesian approach, and the re-parameterization of the model used in the frequentist approach. Nonetheless, Figure B1 shows that the differences between the Bayesian and frequentist approach are small and do not affect the conclusions of this study.

References

Beldring, S., T. Engen-Skaugen, E. J. Forland, and L. A. Roald, 2008: Climate change impacts on hydrological processes in Norway based on two methods for transferring regional climate model results to meteorological station sites, *Tellus A*, **60A**(3), 439–450, doi:10.1111/j.1600-0870.2008.00306.x.

Benestad, R. E., 2010: Downscaling precipitation extremes, *Theor. Appl. Climatol.*, **100**(1-2), 1–21, doi:10.1007/s00704-009-0158-1.

Beven, K., 2009: *Environmental Modelling: An Uncertain Future?*. CRC Press, 328 pp.

Boberg, F., P. Berg, P. Thejll, W. Gutowski, and J. H. Christensen, 2010: Improved confidence in climate change projections of precipitation further evaluated using daily statistics from ensembles models, *Clim. Dynam.*, **35**, 1509–1520, doi:10.1007/s00382-009-0683-8.

Boberg, F., and J. H. Christensen, 2012: Overestimation of Mediterranean summer temperature projections due to model deficiencies, *Nature Climate Change*, **2**(6), 433–436, doi:10.1038/NCLIMATE1454.

Bretherton, C. S., M. Widmann, V. P. Dymnikov, J. M. Wallace, and I. Bladé, 1999: The effective number of spatial degrees of freedom of a time-varying field, *J. Climate*, **12**(7), 1990–2009.

Buser, C. M., H. R. Künsch, D. Lüthi, M. Wild, and C. Schär, 2009: Bayesian multi-model projection of climate: bias assumptions and interannual variability, *Clim. Dynam.*, **33**, 849–868, doi: 10.1007/s00382-009-0588-6.

Buser, C. M., H. R. Künsch., and C. Schär, 2010: Bayesian multi-model projections of climate: generalization and application to ENSEMBLES results, *Clim. Res.*, **44**, 227–241, doi: 10.3354/cr00895.

1 Christensen J. H., T. R. Carter, M. Rummukainen, G. Amanatidis, 2007: Evaluating the
2 performance and utility of regional climate models: the PRUDENCE project, *Climatic Change*, **81**,
3 1–6, doi: 10.1007/s10584-006-9211-6.

4 Christensen, J. H., E. Kjellström, F. Giorgi, G. Lenderink, and M. Rummukainen, 2010: Weight
5 assignment in regional climate models, *Clim. Res.*, **44**(2-3), 179–194, doi:10.3354/cr00916.

6 Déqué, M., D. Rowell, D. Lüthi, F. Giorgi, J. Christensen, B. Rockel, D. Jacob, E. Kjellström, M.
7 de Castro, and B. van den Hurk, 2007: An intercomparison of regional climate simulations for
8 europe: assessing uncertainties in model projections, *Climatic Change*, **81**, 53–70.

9 Dessai, S., and M. Hulme, 2007: Assessing the robustness of adaptation decisions to climate
10 change uncertainties: A case study on water resources management in the east of England, *Global*
11 *Environmental Change*, **17**(1), 59–72.

12 Fischer, A.M., Weigel, A.P., Buser, C.M., Knutti, R., Künsch, H.R., Liniger, M.A., Schär, C., and
13 C. Appenzeller. 2012: Climate change projections for Switzerland based on a Bayesian multi-model
14 approach. *Int. J. Climatol*., **32**, 2348–2371, doi: 10.1002/joc.3396

15 Fowler, A. M., K. J. Hennessy, 1995: Potential impacts of global warming on the frequency and
16 magnitude of heavy precipitation. *Nat. Hazards*, **11**, 283–303.

17 Fowler, H. J., S. Blenkinsop, and C. Tebaldi, 2007: Linking climate change modelling to impacts
18 studies: recent advances in downscaling techniques for hydrological modelling, *Int. J. Climatol.*,
19 **27**(12), 1547–1578.

20 Fowler, H. J. and M. Ekström, 2009: Multi-model ensemble estimates of climate change impacts on
21 UK seasonal precipitation extremes, *Int. J. Climatol.*, **29**(3), 385–416.

Fowler, H.J., M. Ekström, C.G. Kilsby, P.D Jones, 2005: New estimates of future changes in extreme rainfall across the UK using regional climate model integrations 1. Assessment of control climate, *Journal of Hydrology*, **300**(1), 212–233.

Frei, C, R. Schöll, S. Fukutome, J. Schmidli, and P.J. Vidale, 2006: Future change of precipitation extremes in Europe: Intercomparison of scenarios from regional climate models, *Journal of Geophysical Research,* **111**, D06105.

Furrer, R., R. Knutti, S. R. Sain, D. W. Nychka, and G. A. Meeh, 2007: Spatial patterns of probabilistic temperature change projections from a multivariate Bayesian analysis, *Geophys. Res. Lett.*, **34**, L06711.

Gelman, A., J. B. Carlin, H. S. Stern and D. B. Rubin, 2003: *Bayesian Data Analysis*, Chapman & Hall/CRC, 668pp.

Giorgi, F., and L. O. Mearns, 2002: Calculation of average, uncertainty range, and reliability of regional climate changes from AOGCM simulations via the "reliability ensemble averaging" (rea) method, *J. Climate*, **15**, 1141–1158, doi:10.1175/1520-0442.

Hawkins, E., and R. Sutton, 2011: The potential to narrow uncertainty in projections of regional precipitation change, *Clim. Dynam.*, **37**, 407–418.

Haylock, M. R., and C. M. Goodess, 2004: Interannual variability of European extreme winter rainfall and links with mean large-scale circulation, *Int. J. Climatol.*, **24**, 759–776, doi:10.1002/joc.1033.

Herrera, S., L. Fita, J. Fernández, J. M. Gutiérrez, 2010: Evaluation of the mean and extreme precipitation regimes from the ENSEMBLES regional climate multimodel simulations over Spain, Journal of Geophysical Research, **115**(D21), D21117.

IPCC, 2012: Managing the Risks of Extreme Events and Disasters to Advance Climate Change Adaptation. A Special Report of Working Groups I and II of the Intergovernmental Panel on Climate Change [Field, C.B., V. Barros, T.F. Stocker, D. Qin, D.J. Dokken, K.L. Ebi, M.D. Mastrandrea, K.J. Mach, G.-K. Plattner, S.K. Allen, M. Tignor, and P.M. Midgley (eds.)]. Cambridge University Press, Cambridge, UK, and New York, NY, USA, 582 pp.

Kendon, E. J., D. P. Rowell, R. G. Jones, E. Buonomo, 2008: Robustness of future changes in local precipitation extremes, *J. Climate*, **21**(17), 4280–4297.

Knutti, R., R. Furrer, C. Tebaldi, J. Cermak, and G. A. Meehl, 2010: Challenges in combining projections from multiple climate models, *J. Climate*, **23**, 2739–2758, doi:10.1175/2009JCLI3361.1.

Leith, N. A., and R. E. Chandler, 2010: A framework for interpreting climate model outputs, *Appl. Statist.*, **59**(2), 279–296, doi:10.1111/j.1467-9876.2009.00694.x.

Masson, D., R. Knutti, 2011: Climate model genealogy, *Geophys. Res. Lett.*, **38,** L08703, doi:10.1029/2011GL046864.

Maule, C., P. Thejll, J. H. Christensen, S. Svendsen, and J. Hannaford, 2012: Improved confidence in regional climate model simulations of precipitation evaluated using drought statistics from the ensembles models, *Clim. Dynam.*, **40**, 155–173.

Miller R.G., 1974: The jackknife - a review, *Biometrika*, **61**, 1–15

Pennell, C., and T. Reichler, 2011: On the effective number of climate models, *J. Climate*, **24**, 2358–2367, doi:10.1175/2010JCLI3814.1.

Peterson, T. C., 2005: Climate Change Indices, *WMO Bulletin*, **54**(2), 83–86.

1 Räisänen, J., and O. Räty, 2012: Projections of daily mean temperature variability in the future:
2 cross-validation tests with ENSEMBLES regional climate simulations, *Clim. Dynam.*,
3 doi:10.1007/s00382-012-1515-9.

4 Scharling, M., 1999: Klimagrid Danmark nedbør 10∗10 km (ver.2), Technical Report number 99-15
5 available in Danish at www.dmi.dk, Danish Metereological Institute, Denmark.

6 Sibson, R., 1980: A vector identity for the dirichlet tessellation, *Math. Proc. Cambridge*, **87**(1),
7 151–155, doi:10.1017/S0305004100056589.

8 Sibson R., 1981: *Interpreting Multivariate Data*, Wiley, 374.

9 Sunyer, M. A., H. Madsen, D. Rosbjerg, and K. Arnbjerg-Nielsen, 2013a: Regional
10 interdependency of precipitation indices across Denmark in two ensembles of high resolution
11 RCMs, *J. Climate*, **26**, 7912–7928, doi:10.1175/JCLI-D-12-00707.1.

12 Sunyer, M. A., H. J. D. Sørup, O. B. Christensen, H. Madsen, D. Rosbjerg, P. S. Mikkelsen, and K.
13 Arnbjerg-Nielsen, 2013b: On the importance of observational data properties when assessing
14 regional climate model performance of extreme precipitation, *Hydrol. Earth Syst. Sci. Discuss.*, **10**,
15 1–41, doi:10.5194/hessd-10-1-2013.

16 Taylor, K. E., J. R. Stouffer, and G. A. Meehl, 2012: An Overview of CMIP5 and the Experiment
17 Design. *Bull. Amer. Meteor. Soc.*, **93**, 485–498, doi: http://dx.doi.org/10.1175/BAMS-D-11-
18 00094.1

19 Tebaldi, C., L. O. Mearns, D. Nychka, and R. L. Smith, 2004: Regional probabilities of
20 precipitation change: a bayesian analysis of multimodel simulations, *Geophys. Res. Lett.*, **31**,
21 L24213.

Tebaldi, C., R. Smith, D. Nychka, and L. O. Mearns, 2005: Quantifying uncertainty in projections of regional climate change: a Bayesian approach to the analysis of multi-model ensembles, *J. Climate*, **18**, 1524–1540, doi:10.1175/JCLI3363.1.

Tebaldi, C., and R. Knutti, 2007: The use of the multi-model ensemble in probabilistic climate projections, *Philos. Trans. Roy. Soc. A*, **365**, 2053–2075.

van der Linden, P. and J. F. Mitchell, 2009: Ensembles: Climate change and its impacts: Summary of research and results from the ensembles project, Technical Report, Met Office Hadley Centre, Exeter, UK.

Weigel, A.P., R. Knutti, M.A. Liniger, and C. Appenzeller 2010: Risks of model-weighting in multimodel climate projections, *J. Clim*, **23**, 4175-4191.

Wilby, R. L., and I. Harris, 2006: A framework for assessing uncertainties in climate change impacts: Low-flow scenarios for the river thames, UK, *Water Resour. Res.*, **42**(2), 1–10.

Wilks, D. S., 2006: *Statistical Methods in the Atmospheric Sciences, second edition*. Academic Press, 627 pp.

1 Table 1 - List of RCMs used in this study, driving GCMs, and source of the RCMs.

RCM	**GCM**	**Institute**
HIRHAM5	ARPEGE	Danish Meteorological Institute
HIRHAM5	ECHAM5	
HIRHAM5	BCM	
REMO	ECHAM5	Max Planck Institute for Meteorology
RACMO2	ECHAM5	Royal Netherlands Meteorological Institute
RCA	ECHAM5	Swedish Meteorological and Hydrological Institute
RCA	BCM	
RCA	HadCM3Q3	
CLM	HadCM3Q0	Swiss Federal Institute of Technology, Zürich
HadRM3Q0	HadCM3Q0	UK Met Office
HadRM3Q3	HadCM3Q3	
HadRM3Q16	HadCM3Q16	
RCA3	HadCM3Q16	Community Climate Change Consortium for Ireland
RM5.1	ARPEGE	National Centre for Meteorological Research in France
RegCM3	ECHAM5	International Centre for Theoretical Physics

2

3

Table 2 Prior distributions and hyperparameters. In the case of normal distribution the parameters are the mean and variance. In the case of gamma distribution the parameters are the shape and scale.

Parameter	**Distribution**	**Mean/Shape**	**Variance/Scale**
μ	Normal	X_{Obs}	1000
v	Normal	$X_{obs} + (\overline{Y} - \overline{X})/(A+1)$	1000
β	Normal	$\overline{X} - X_{obs}$	1000
λ	Gamma	0.001	1/0.001
θ	Gamma	0.001	1/0.001

Figure caption list

Figure 1 – Linear regressions between the bias, $ABias_{..}$, and observations, $X_{Obs,.}$, for winter, summer and annual period. The grey lines show the linear regressions fitted to each RCM individually, the black lines show the linear regressions fitted considering the biases of all the RCMs together. The dots show the values estimated for all grid points and all RCMs. The figure illustrates how α is estimated from the values of μ, v, and β.

Figure 2 – **X** (grey points) and **Y** (black points) vs. theoretical values of a normal distribution for winter, summer and annual period. The thick lines show the identity line. The thin lines represent the 95% confidence limits.

Figure 3 – Figure 3 –Posterior (black line in primary y-axis) and prior (grey shaded area in secondary y-axis) probability density functions of the five parameters in the model (μ, v, β, λ, and θ) for winter, summer and annual period. The grey dots are the outputs from each of the RCMs and the black dots are the values estimated from the observations. The vertical thick dashed lines represent the median, while the thin lines represent the 5^{th} and 95^{th} percentiles. The black and grey solid lines represent the mean obtained using the Bayesian and frequentist approach, respectively.

Figure 4 - Scatter plots for all the pairs of parameters and histograms of the marginal posterior distributions of each parameter for the summer period. In the diagonal from top left to bottom right: μ [mm day^{-1}], v [mm day^{-1}], β [mm day^{-1}], λ [mm^{-2} day^{2}], and θ [-]. The values of RRwn95 obtained from the RCMs for present and future period are also shown in the scatter plot for the pair of parameters [μ, v].

Figure 5 – Marginal posterior distributions of μ, v, and CF for three different tests. The grey dots are the outputs from each of the RCMs and the black dots are the values estimated from the observations. The vertical thick line represents the median, while the thin lines represent the 5^{th} and

95th percentiles. The dotted line indicates CF equal to 1, while the dash-dotted line shows the median of the CFs estimated from the RCMs in the ensemble.

Figure 6 – Scatter plot for μ and v for three different tests for the winter, summer, and annual period.

Figure 7 – Median and standard deviation of CF for a set of tests. The y- and x-axes show the factors that are multiplied to **R** and A, respectively. The colour of the circles indicates the distribution median and the size of the circles indicates the standard deviation. Larger circle sizes correspond to higher standard deviation. The number 1, 2, and 3 indicate the three tests analysed in detail

Figure B1 – Mean and standard error of the parameters in Eq. (2) and CF estimated using a Bayesian (black) and frequentist approach (grey) for the three tests described in section 4.c. The circles indicate the mean and the lines represent the standard error (the total length of the line is equal to the standard error).

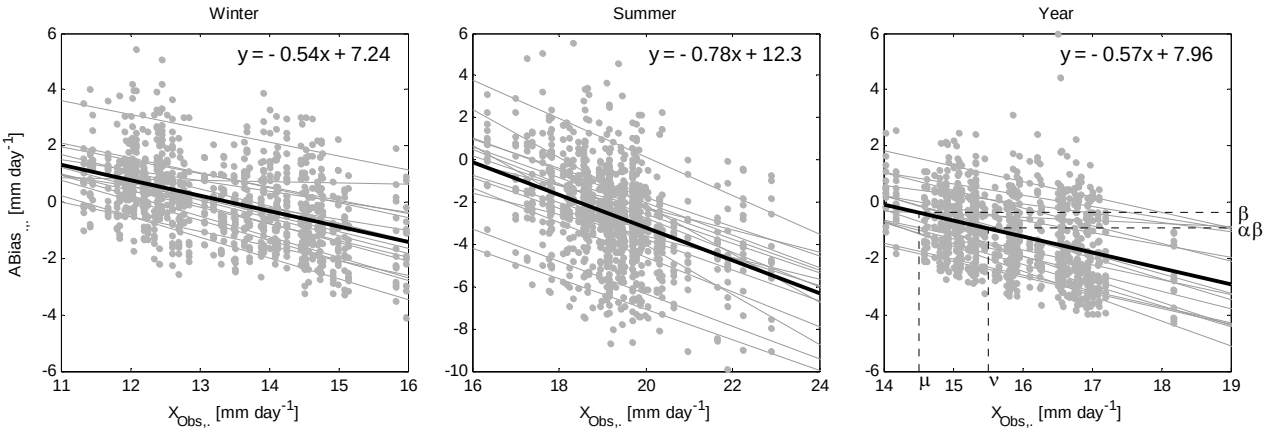

Figure 1 – Linear regressions between the bias, $ABias_{.,.}$, and observations, $X_{Obs,.}$, for winter, summer and annual period. The grey lines show the linear regressions fitted to each RCM individually, the black lines show the linear regressions fitted considering the biases of all the RCMs together. The dots show the values estimated for all grid points and all RCMs. The figure illustrates how α is estimated from the values of μ, v, and β.

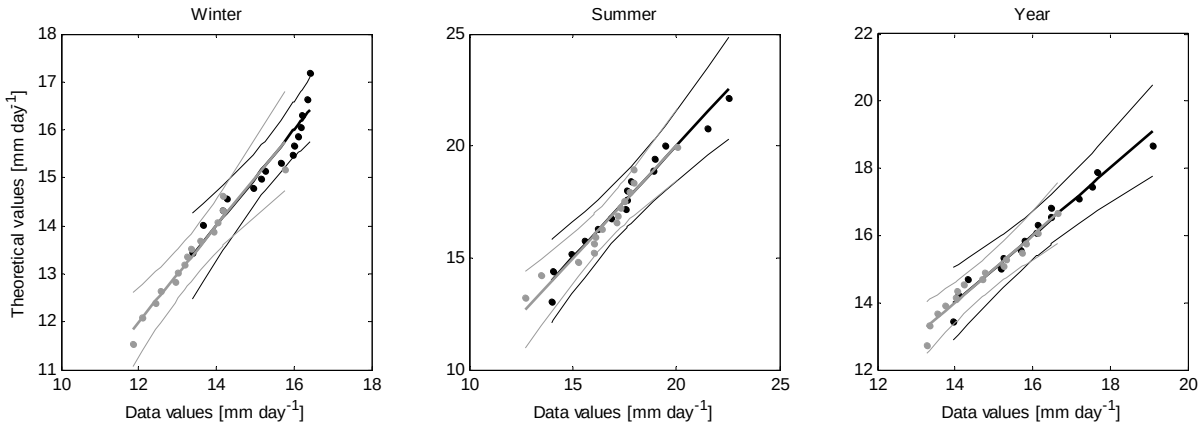

Figure 2 – **X** (grey points) and **Y** (black points) vs. theoretical values of a normal distribution for winter, summer and annual period. The thick lines show the identity line. The thin lines represent the 95% confidence limits.

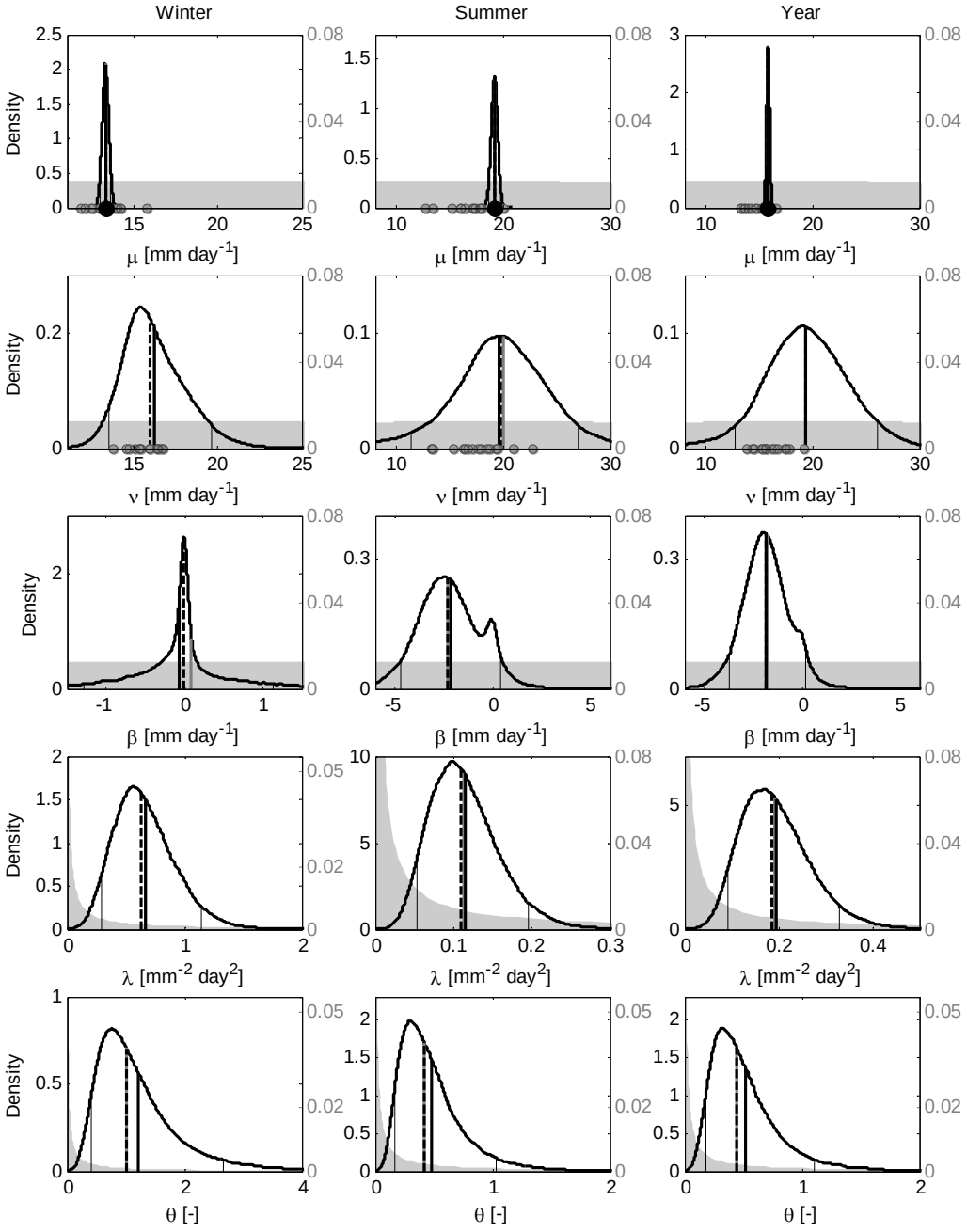

Figure 3 –Posterior (black line in primary y-axis) and prior (grey shaded area in secondary y-axis) probability density functions of the five parameters in the model (μ, v, β, λ, and θ) for winter, summer and annual period. The grey dots are the outputs from each of the RCMs and the black dots are the values estimated from the observations. The vertical thick dashed lines represent the median, while the thin lines represent the 5th and 95th percentiles. The black and grey solid lines represent the mean obtained using the Bayesian and frequentist approach, respectively.

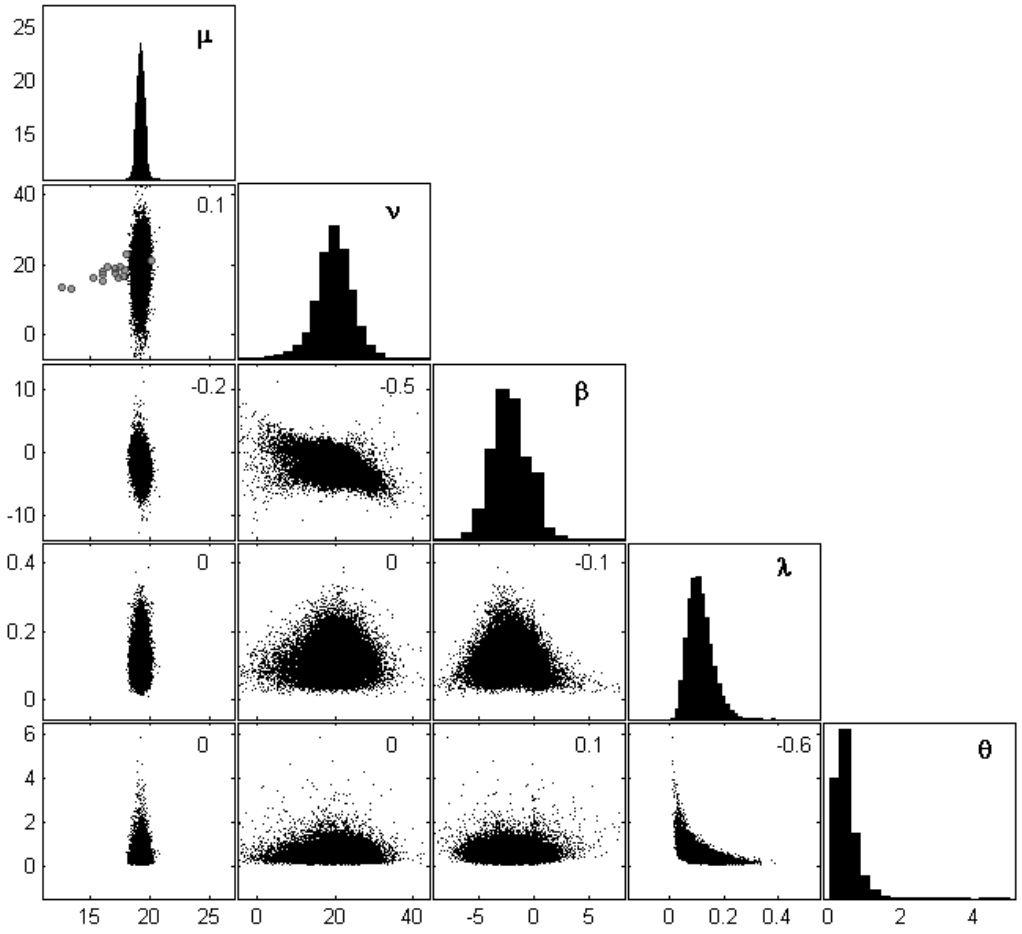

Figure 4 - Scatter plots for all the pairs of parameters and histograms of the marginal posterior distributions of each parameter for the summer period. In the diagonal from top left to bottom right: μ [mm day^{-1}], v [mm day^{-1}], β [mm day^{-1}], λ [mm^{-2} day^{2}], and θ [-]. The values of RRwn95 obtained from the RCMs for present and future period are also shown in the scatter plot for the pair of parameters [μ, v]. The number shown in the plots is the linear correlation coefficient.

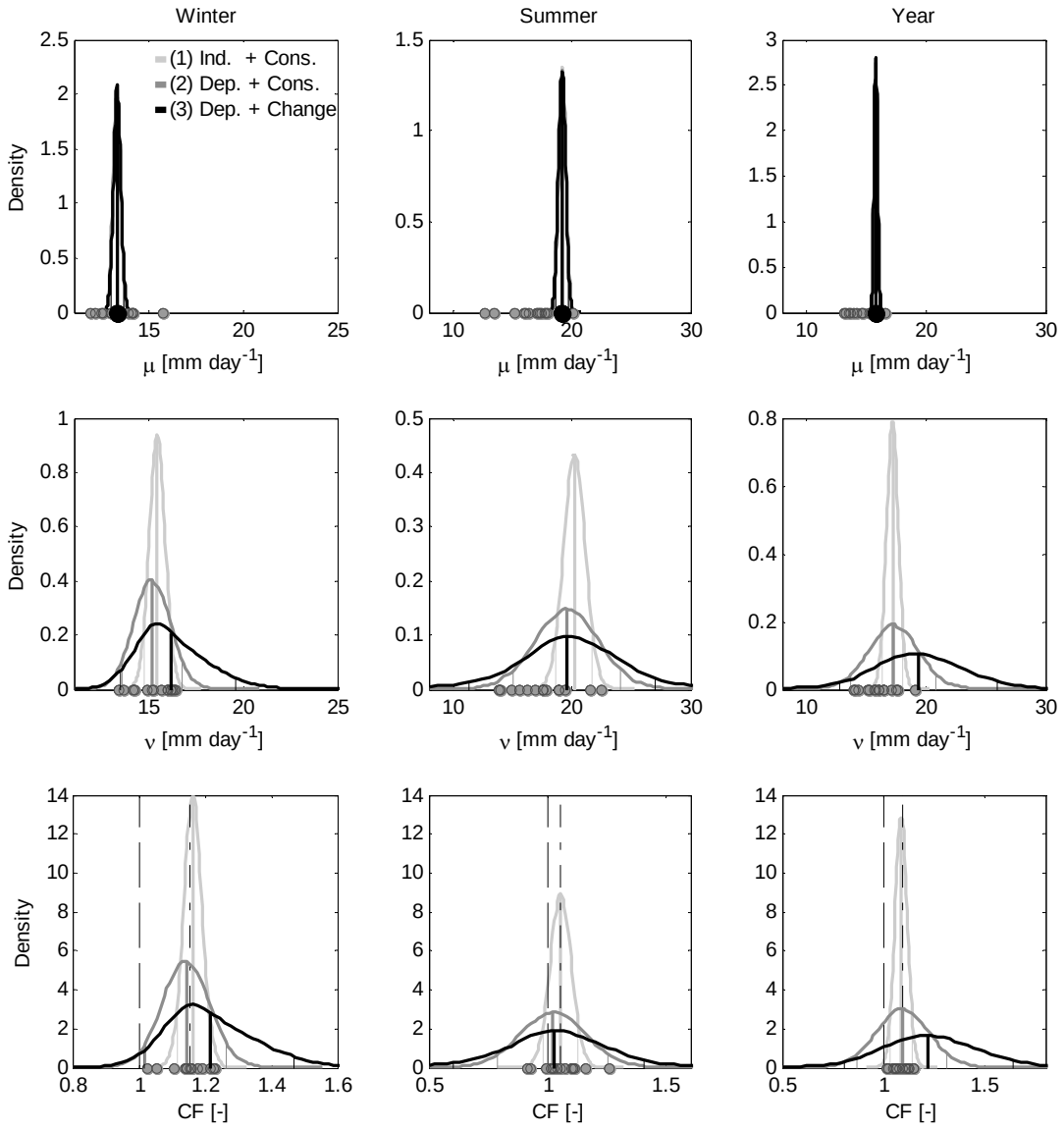

Figure 5 – Marginal posterior distributions of μ, ν, and CF for three different tests. The grey dots are the outputs from each of the RCMs and the black dots are the values estimated from the observations. The vertical thick line represents the mean, while the thin lines represent the 5th and 95th percentiles. The dashed line indicates CF equal to 1, while the dash-dotted line shows the median of the CFs estimated from the RCMs in the ensemble.

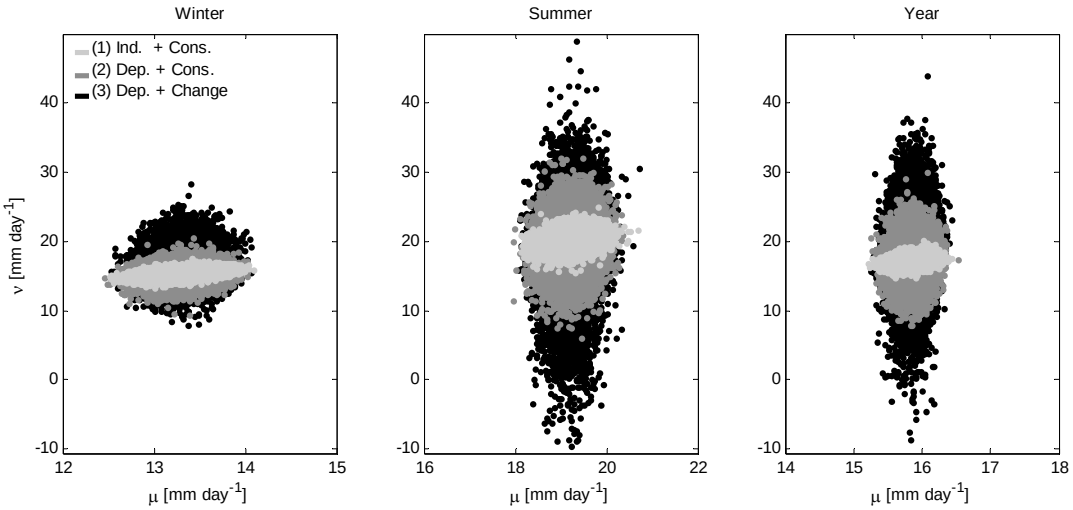

Figure 6 – Scatter plot for μ and v for three different tests for the winter, summer, and annual period.

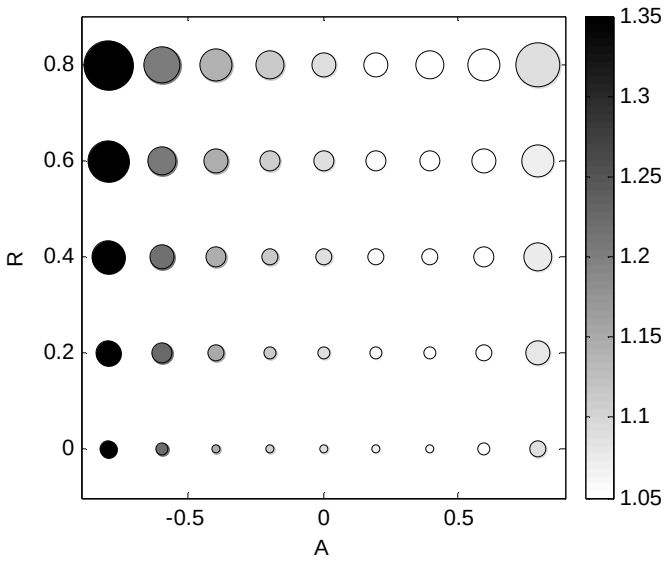

Figure 7 – Mean and standard deviation of CF for a set of tests. The y- and x-axes show the values used to define the correlation coefficients in **R** and A, respectively. The colour of the circles indicates the distribution mean and the size of the circles indicates the standard deviation. Larger circle sizes correspond to higher standard deviation.

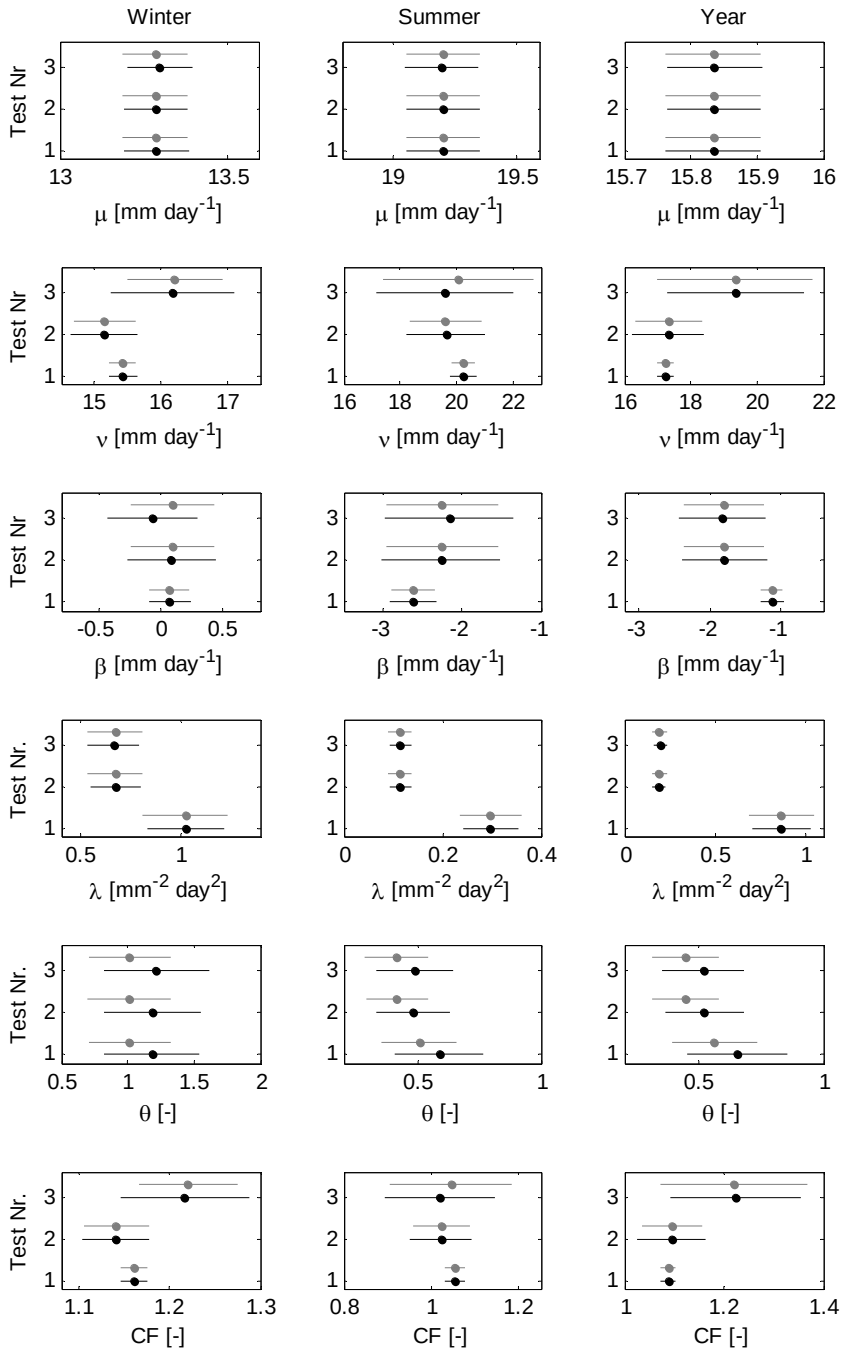

Figure B1 – Mean and standard error of the parameters in Eq. (2) and CF estimated using a Bayesian (black) and frequentist approach (grey) for the three tests described in section 4.c. The circles indicate the mean and the lines represent the standard error (the total length of the line is equal to the standard error).